U. S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Weather Service

NOAA Technical Memorandum NWS SR-62

AN ALTERNATE TO THE HURRAN (Hurricane Analog) TROPICAL CYCLONE FORECAST SYSTEM

SOUTHERN REGION HEADQUARTERS
SCIENTIFIC SERVICES DIVISION
FORT WORTH, TEXAS
January 1972

AN ALTERNATE TO THE HURRAN (Hurricane Analog) TROPICAL CYCLONE FORECAST SYSTEM

Charles J. Neumann, National Hurricane Center, Miami, Florida

ABSTRACT

The HURRAN (Hurricane Analog) system is one of several techniques from which the National Hurricane Center derives objective guidance preparatory to the issuance of tropical cyclone track forecasts. One shortcoming of the HURRAN system is its inability to provide objective guidance when insufficient analogs are found. Accordingly, an alternate regression equation system known as CLIPER (Climatology and Persistence) was developed. This paper describes the derivation, application and errors of the CLIPER system.

INTRODUCTION

The National Hurricane Center (NHC) currently uses three semi-independent computerized techniques to provide objective guidance preparatory to the issuance of tropical cyclone track forecasts. These systems, known as NHC-67, HURRAN and SANBAR, provide estimates of storm displacement over forecast periods extending to 72 hours.

The NHC-67 system (Miller, et al, 1968) and its predecessor, NHC-64 (Miller and Chase, 1966) has been in use at NHC for a number of years. The system computes storm displacement from a series of regression equations using predictors derived from the observed heights of the 1000, 700 and 500-mb surfaces.

HURRAN (Hope and Neumann, 1970) is an analog system. The recorded tropical cyclone tracks back to the year 1886 are computer-scanned and those with time and space characteristics similar to the current storm are identified and displaced to a common origin. The cluster of analog storm positions at the various forecast intervals are then fitted to a bivariate normal distribution, the centroids of these distributions representing the forecast track. Both HURRAN and NHC-67 use some persistence in the early forecast periods.

SANBAR (Sanders and Burpee, 1968) is a filtered barotropic model with input derived from grid-point values of the current 1000 to 100-mb pressure weighted winds. Some "bogus" data are required to augment the wind field in sparse data regions.

(This page intentionally left blank)

These three systems represent entirely different approaches to the problem of tropical cyclone forecasting, and each is capable of producing displacement forecasts with vector errors ranging from near zero to over 1000 n.mi. for the 72-hr. forecast period. It is not unusual for the systems to follow widely different forecast tracks. Under such conditions, it is difficult for the hurricane forecaster to make a decision as to which forecast track is apt to have minimum error. Although objective guidelines are currently being used (Simpson, 1971), the problem of which forecast track to follow, if any, remains as one of the critical decisions for the hurricane forecaster.

PURPOSE

One obvious solution to the problem outlined in the preceding section is to combine the good features of each of the two statistical schemes, NHC-67 and HURRAN, into a single system. Such an approach led to the original development of the NHC-64 system and is an approach currently being taken (among others) by researchers at the National Hurricane Center.

It is known, for example, that HURRAN does well with tropical cyclones exhibiting "normal" tracks. For the most part, storms south of 25° N., before recurvature, behave normally and are forecast well by the HURRAN system. A detailed error analysis of HURRAN (Neumann and Hope, 1971) shows that the system does not do as well farther north, after storm recurvature. NHC-67, on the other hand, because of its synoptic data input (completely lacking in HURRAN) does relatively better than HURRAN with increasing latitude.

One of the stated shortcomings of the HURRAN system is its inability to find sufficient analogs for a track forecast in about one out of three storms. This would prove to be a serious difficulty when one attempts to statistically combine HURRAN with some other system. Accordingly, an alternate system, known as CLIPER (climatology and persistence) was devised. The system uses a series of regression equations fitted to essentially the same predictors used in the analog sense by HURRAN. The primary purpose of this study is to discuss details on the derivation of the CLIPER system and to present a preliminary error analysis based on the dependent data set. Conceptually, CLIPER was intended as a back-up for HURRAN. However, as will be pointed out in a subsequent section, CLIPER appears to outperform HURRAN in some areas.

THE DEVELOPMENT DATA SET

The basic input to both the HURRAN and the CLIPER systems are historical storm track data residing on a magnetic tape maintained and continually updated at the National Hurricane Center. These data were originally compiled at the National Weather Records Center from data given by Cry, et al, 1959. Further details on the data set are given by Hope and Neumann, 1968.

Both the HURRAN and CLIPER systems mathematically re-create the original storm tracks from the digitized storm positions given on the data tape. This is accomplished by locally fitting a third-order polynomial to each of four time-ordered storm positions with some additional smoothing to insure slope-continuity when moving to the next polynomial. A somewhat similar fitting process is discussed by Akima, 1970. The HURRAN program reads the historical data from either tape or disk each time the program is executed. In the CLIPER system, the tape is read only once and a series of multiple regression equations are fitted to essentially the same predictors incorporated in HURRAN. Additionaly, CLIPER uses storm intensity as a predictor. This further stratification contributes to a decrease in the standard error of the CLIPER dependent data compared to the HURRAN dependent data.

In order to insure a homogeneous data set for each forecast period, 12 through 72 hours, all storms with recorded life histories of less than five days were eliminated from the dependent data set. This insured that for each storm there were at least three sets of 12 through 72-hour forecast positions and a previous 12-hour position from which to derive persistence. All storms prior to 1931 were excluded from the data, leaving a total of 3156 sets of dependent data on 286 storms over the 40-year period 1931 through 1970. Figure 1 shows the geographical bounds of the CLIPER dependent data sample. As will be discussed later, one must use caution when applying regression equations outside the time and space bounds of the dependent data sample.

THE REGRESSION ANALYSIS

For each of the 3156 sets of dependent data, eight basic predictors were available. These, together with their means and standard deviations are given in Table 1. The twelve predictands together with their means and standard deviations are given in Table 2.

Table 1. Means and Standard Deviations of the Eight Basic Predictors

Predictor	Symbol	Mean	Standard Deviation
Initial longitude	X_0	68.4	15.4 degs.
Initial latitude	Y_0	24.1	7.3 degs.
Initial E. to W. component	U_0	-3.4	8.8 kts.
E. to W. component 12 hrs. ago	U_{-12}	-4.2	8.5 kts.
Initial S. to N. component	V_0	5.1	5.2 kts.
S. to N. component 12 hrs. ago	V_{-12}	4.9	4.9 kts.
Maximum wind	W	71.4	32.7 m.p.h.
Day number	D	247.9	37.0 ---

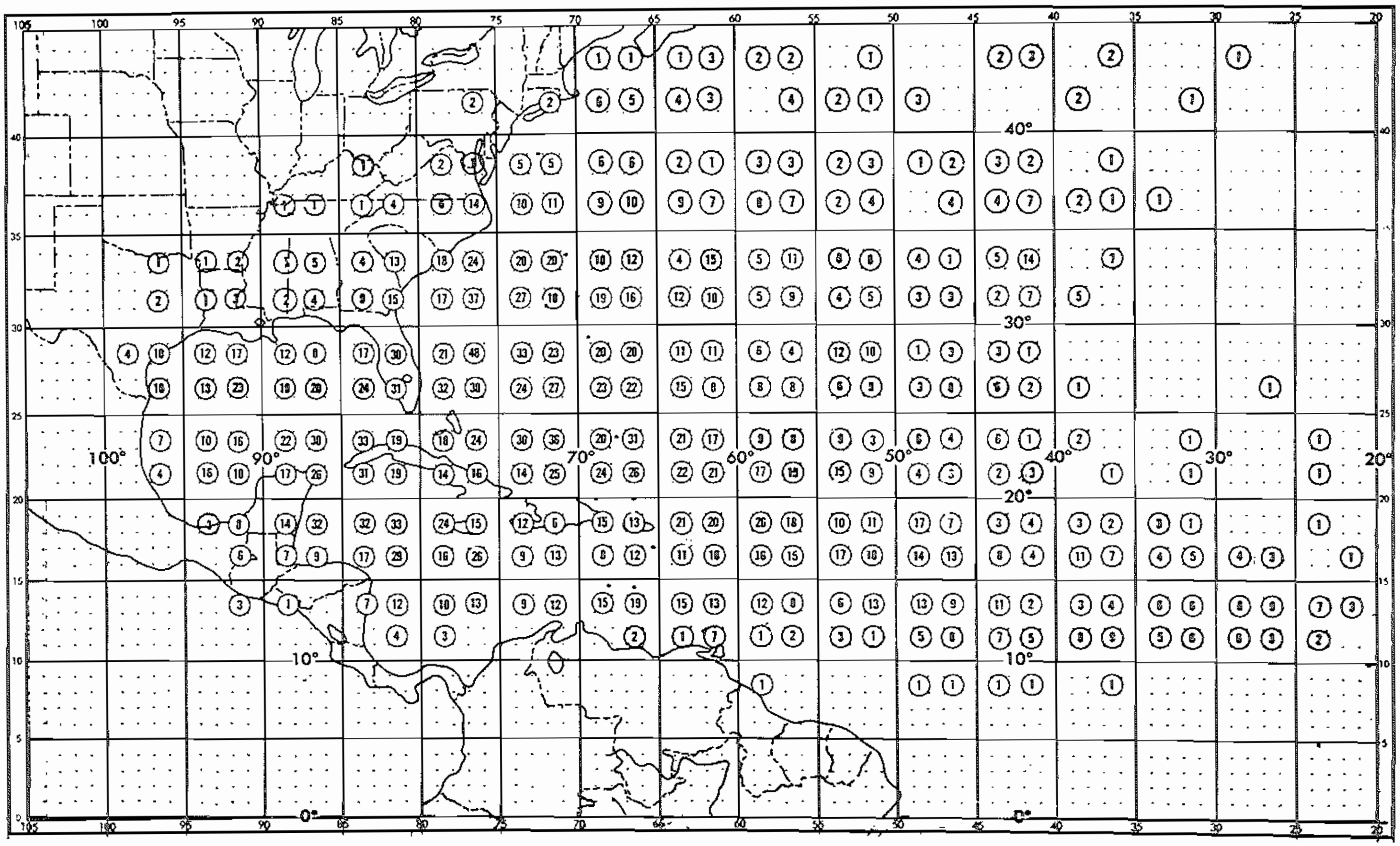

Figure 1. Geographical distribution of the 3,156 cases comprising the CLIPER data set. The circled numbers give the count of initial storm positions located within each 2½ deg. latitude/longitude box.

Table 2. Means and Standard Deviations (Nautical Miles) of the Twelve Predictands

Predictand	Symbol	Mean	Standard Deviation
12 hr. N/S displacement*	DY_{12}	62	64
24 hr. N/S displacement	DY_{24}	128	124
36 hr. N/S displacement	DY_{36}	197	185
48 hr. N/S displacement	DY_{48}	270	245
60 hr. N/S displacement	DY_{60}	347	307
72 hr. N/S displacement	DY_{72}	429	372
12 hr. E/W displacement	DX_{12}	-35	107
24 hr. E/W displacement	DX_{24}	-59	215
36 hr. E/W displacement	DX_{36}	-70	325
48 hr. E/W displacement	DX_{48}	-68	438
60 hr. E/W displacement	DX_{60}	-53	552
72 hr. E/W displacement	DX_{72}	-26	669

*Southward and westward motion is negative.

The general form of the regression equation used here where DISP refers to any given displacement in either the meridional (y) or the zonal (x) direction is given by,

$$DISP = f(X_0, Y_0, U_0, U_{-12}, V_0, V_{-12}, W, D) \tag{1}$$

where the symbols have the same meaning as given in Table 1. If only the observed values (as opposed to derived quantities or higher order terms) of the predictors are used in the analysis then (1) takes the familiar form,

$$DISP = C_0+C_1X_0+C_2Y_0+C_3U_0+C_4U_{-12}+C_5V_0+C_6V_{-12}+C_7W+C_8D \tag{2}$$

where the one-dimensional array C are constants. Equation (2) assumes a linear relationship between DISP and each of the predictors. A better estimate of DISP might be obtained by considering a higher order predictor function. For illustration, let DISP be a function of only one predictor R,

$$DISP = f(R). \tag{3}$$

A non-linear estimate of DISP can be provided by the third order polynomial,

$$DISP = f(R) = C_0+C_1R+C_2R^2+C_3R^3. \tag{4}$$

Similarly, if DISP is a function of two predictors R and Q, a non-linear estimate of DISP may be obtained from,

$$DISP = f(R,Q) = C_0+C_1R+C_2R^2+C_3R^3+C_4Q+C_5QR+C_6QR^2+C_7Q^2+C_8Q^2R+C_9Q^3. \tag{5}$$

Panofsky, 1949, used polynomial (5) as the basis of an objective analysis scheme relating 500mb height to an x-y coordinate system.

Note that the number of constants (and terms) on the right side of (4) and (5) increased from 4 for one predictor to 10 for 2 predictors. As additional predictors (P) are added, the number of terms (T) increases according to the number of combinations of P+3 distinct predictors taken 3 at a time and is given by,

$$T = (P+3)!/6(P!). \tag{6}$$

Since one of the terms always includes an intercept value, the number of predictors is given by T-1. If all 8 basic predictors are included in the third-order polynomial representation of (1) there are a total of 165 possible combinations involving the products or cross-products of the original 8 predictors. The individual terms can be obtained from,

$$DISP(P_1,P_2\ldots P_8) = \sum_{j_1\ldots j_8}\left(C_{j_1\ldots j_8}\prod_{i=1}^{8} P_i^{j_i}\right) \qquad (j_1+j_2+\ldots j_8 \leq 3) \tag{7}$$

where array C are constants.

PREDICTOR SELECTION

In the initial stepwise multiple screening regression run of the type described in Efroymson, 1964, the analysis was continued through all of the 164 predictors given by (7) regardless of the reduction of variance provided by each predictor. Even with all 164 predictors in the regression equation, F-test criteria (Burington and May, 1958) still showed statistical significance at the 1 percent level, the loss of the degrees of freedom being compensated by the large number of cases. The plot marked "A" on figure 2 shows the reduction of variance provided by fitting (7) to each of the 12 predictands where the reduction in variance (RV) is given by,

$$RV = R_m^2 = 1.0-(SE^2/SD^2) \tag{8}$$

where R_m is the multiple correlation coefficient, SE is the standard deviation of the predictand about the regression function (standard error) and SD is the standard deviation of the predictand about its mean.

In order to assess the partial contribution of smaller groups of variables, the regression analysis was also carried out by forcing out of the regression all terms other than those involving initial motion and day number (plot B on figure 2) and again by forcing out all terms other than those involving initial position and day number (plot C on figure 2). As pointed out by Mills, 1955, such a procedure of eliminating the effects of some of the variables can provide considerable diagnostic information.

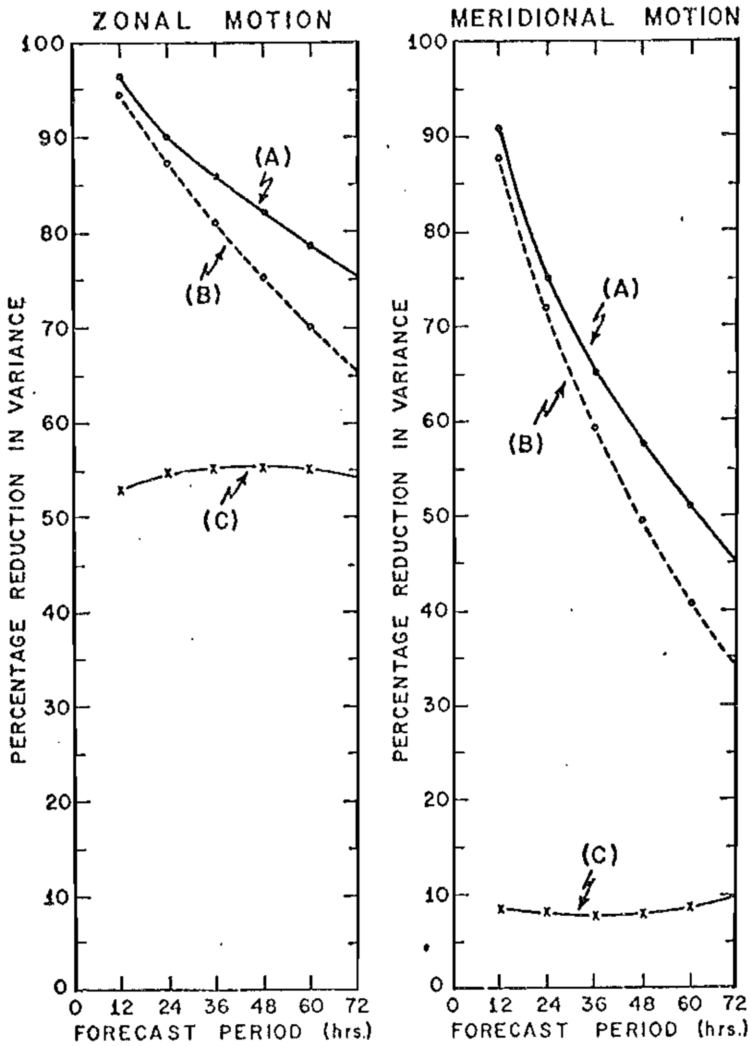

Figure 2. Reduction of variance (%) vs. forecast period using A, all 164 predictors; B, the 19 predictors involving initial motion and day number, and C, the 19 predictors involving initial position and day number.

In the broad sense, figure 2 points out that the meridional motion is statistically less predictable than is the zonal motion. This is particularly true in the extended forecast periods. This strongly suggests that variations in the environmental steering mechanism, not accounted for in the CLIPER system, affect the meridional motion more than the zonal motion. Accordingly, the addition of synoptic predictors to CLIPER (or HURRAN), to be reported on in a separate paper, should be expected to improve the meridional forecast more than the zonal forecast.

The reductions of variance shown in figure 2 are the greatest which can be realized by knowing the precise values of the appropriate basic predictors listed in Table 1. Obviously, the total reduction would be less if the input data were independent rather than dependent. The relatively high reductions of variance provided by the persistence and day number function (plots B) point out the extreme importance of knowing as precisely as possible the initial and past motion of a storm. In this connection, Hope, 1971, presents data showing that the accuracy of storm center location is a direct function of the number of aircraft fixes.

Plots C on figure 2 show the reductions in variance provided if only the initial position and day number are considered in the analysis. This information appears to be moderately useful as a tool in reducing the variance in zonal motion but is practically useless as a tool in predicting meridional motion. That is to say, there is apparently no statistical way of improving on a forecast of meridional storm displacement by simply knowing the latitude, longitude, and time of year.

ELIMINATION OF NON-SIGNIFICANT PREDICTORS

Many of the 165 terms generated by (7) are highly inter-correlated and do not significantly reduce the variance. Subsequent computer runs were programmed to terminate the regression analysis when a total of 9 predictors from the original 164 had been selected for inclusion in the multiple regression equation. This was usually comparable to terminating the analysis when an additional predictor failed to lower the variance an additional one percent. Tables 3 and 4 show the resultant predictor selection order and the associated reduction of variance for both zonal and meridional motion. As expected, persistence, based on

Table 3. Predictor Selection Order and Reduction of Variance (Percent) for East/West Displacement

Predictor	Predictand					
	DX_{12}	DX_{24}	DX_{36}	DX_{48}	DX_{60}	DX_{72}
U_0	1(94.3)	1.(84.9)	1(76.8)	1(68.6)	1(61.4)	1(55.1)
U_{-12}	2 (0.9)					
Y_0		2 (0.9)	2 (2.1)	2 (3.7)	2 (5.3)	2 (6.8)
V_0		3 (0.8)	3 (1.7)	3 (2.8)	3 (3.8)	3 (4.5)
$V_0^2U_{-12}$		4 (0.6)				
$Y_0V_0U_{-12}$			4 (1.1)	4 (1.7)	4 (2.1)	4 (2.4)
X_0						5 (0.5)
Total Reduction	(95.1)	(87.2)	(82.1)	(76.9)	(72.6)	(69.4)

Table 4. Predictor Selection Order and Reduction of Variance (Percent) for North/South Displacement

Predictor	Predictand					
	DY_{12}	DY_{24}	DY_{36}	DY_{48}	DY_{60}	DY_{72}
V_0	1(87.4)	1(69.4)	1(57.4)	1(46.3)	1(37.0)	1(29.4)
V_{-12}	2 (1.7)					
$V_0V_{-12}^2$		2 (0.6)				
WV_{-12}		3 (0.5)				
V_0W			2 (0.9)	2 (1.4)	2 (1.8)	2 (2.0)
$V_0^2V_{-12}$			3 (1.1)			
$Y_0^2V_0$				3 (1.5)	3 (2.0)	3 (2.3)
D^2V_{-12}				4 (0.5)	4 (0.7)	4 (0.7)
U_0D^2				5 (0.6)	5 (0.8)	5 (0.9)
Y_0^2D						6 (0.5)
WDV_{-12}						7 (0.5)
U_0						8 (0.5)
D^2						9 (0.5)
Total Reduction	(89.0)	(70.6)	(59.4)	(50.3)	(42.2)	(36.9)

current storm motion was selected as the prime predictor for each of the 12 displacement forecasts. Latitude (Y_0) was selected as a secondary predictor on most of the zonal displacement forecasts but neither latitude nor longitude (except in combination with other predictors) was selected as significant in the case of the meridonial motion. Also significant is the fact that all eight predictors or functions thereof were selected for meridional displacement but only initial position and motion were selected in the case of the zonal displacement.

It is difficult to physically reason why some of the higher order terms were selected for retention in the prediction equations and others were not. Although the correlation coefficient matrix is available for study and is included as Tables 5, 6, and 7, complex inter-correlations between predictors and predictor functions make such an analysis

Table 5. Linear Correlation Coefficients Between Basic Predictors and Predictands

	U_0	V_0	U_{-12}	V_{-12}	X_0	Y_0	W	D
DY_{12}	0.25	0.93	0.16	0.65	0.04	0.20	0.13	-0.03
DY_{24}	0.23	0.83	0.14	0.61	0.03	0.18	0.16	-0.03
DY_{36}	0.21	0.76	0.13	0.55	0.03	0.17	0.18	-0.02
DY_{48}	0.20	0.68	0.12	0.49	0.02	0.16	0.19	-0.02
DY_{60}	0.19	0.61	0.11	0.43	0.02	0.15	0.20	-0.01
DY_{72}	0.18	0.54	0.11	0.38	0.01	0.16	0.21	-0.01
DX_{12}	0.94	0.30	0.84	0.32	0.21	0.65	0.05	0.07
DX_{24}	0.92	0.33	0.81	0.35	0.23	0.66	0.06	0.08
DX_{36}	0.88	0.37	0.76	0.38	0.23	0.67	0.07	0.08
DX_{48}	0.83	0.40	0.72	0.40	0.24	0.68	0.09	0.08
DX_{60}	0.78	0.41	0.68	0.40	0.24	0.68	0.11	0.08
DX_{72}	0.74	0.42	0.65	0.41	0.24	0.68	0.12	0.08

Table 6. Linear Correlations Between the Basic Predictors

	U_0	V_0	U_{-12}	V_{-12}	X_0	Y_0	W	D
U_0	1.00	0.26	0.91	0.29	0.19	0.64	0.04	0.07
V_0		1.00	0.19	0.78	0.04	0.24	0.12	-0.03
U_{-12}			1.00	0.23	0.17	0.62	0.02	0.07
V_{-12}				1.00	0.05	0.30	0.09	-0.05
X_0					1.00	0.01	0.03	-0.12
Y_0						1.00	0.19	0.02
W							1.00	0.12
D								1.00

Table 7. Linear Correlations Between Predictands

	DY_{12}	DY_{24}	DY_{36}	DY_{48}	DY_{60}	DY_{72}	DX_{12}	DX_{24}	DX_{36}	DX_{48}	DX_{60}	DX_{72}
DY_{12}	1.0	.94	.87	.79	.71	.64	.29	.33	.36	.40	.42	.43
DY_{24}		1.0	.96	.90	.82	.75	.46	.33	.37	.41	.44	.46
DY_{36}			1.0	.97	.91	.84	.26	.32	.37	.41	.45	.47
DY_{48}				1.0	.98	.92	.25	.30	.36	.41	.45	.48
DY_{60}					1.0	.98	.23	.29	.35	.40	.45	.48
DY_{72}						1.0	.22	.28	.34	.39	.44	.47
DX_{12}							1.0	.97	.93	.89	.84	.80
DX_{24}								1.0	.98	.95	.91	.87
DX_{36}									1.0	.99	.96	.92
DX_{48}										1.0	.99	.96
DX_{60}											1.0	.99
DX_{72}												1.0

extremely difficult. Some insight into the complexities may be realized by considering Table 8. The table shows that neither day number, day number squared, nor day number cubed, considered singly, correlates significantly* with storm displacement. However, by taking these terms in combination in polynomial form, a much better and statistically significant multiple correlation coefficient is obtained. The solution to the 24-hr fitted polynomial is shown in figure 3.

Table 8. Correlation Coefficients Between Predictands and Selected Functions of Day Number

	F1(D)	F2(D)	F3(D)	F4(D)	F5(D)
DY_{12}	-0.03	-0.05	0.04	0.09	0.09
DY_{24}	-0.03	-0.07	0.04	0.10	0.10
DY_{36}	-0.02	-0.08	0.05	0.11	0.11
DY_{48}	-0.02	-0.09	0.05	0.12	0.12
DY_{60}	-0.01	-0.10	0.06	0.13	0.13
DY_{72}	-0.01	-0.11	0.06	0.14	0.14
DX_{12}	0.07	0.12	-0.02	0.21	0.23
DX_{24}	0.08	0.11	-0.02	0.20	0.22
DX_{36}	0.08	0.11	-0.02	0.19	0.21
DX_{48}	0.08	0.10	-0.02	0.18	0.20
DX_{60}	0.08	0.09	-0.01	0.18	0.20
DX_{72}	0.08	0.09	-0.01	0.17	0.19

$F1(D) = D$

$F2(D) = D^2$

$F3(D) = D^3$

$F4(D) = C_1 + C_2(D) + C_3(D)^2$

$F5(D) = C_1 + C_2(D) + C_3(D)^2 + C_4(D)^3$

*Considering the 3156 cases in the data set, statistical significance starts with correlation coefficients between plus or minus 0.05.

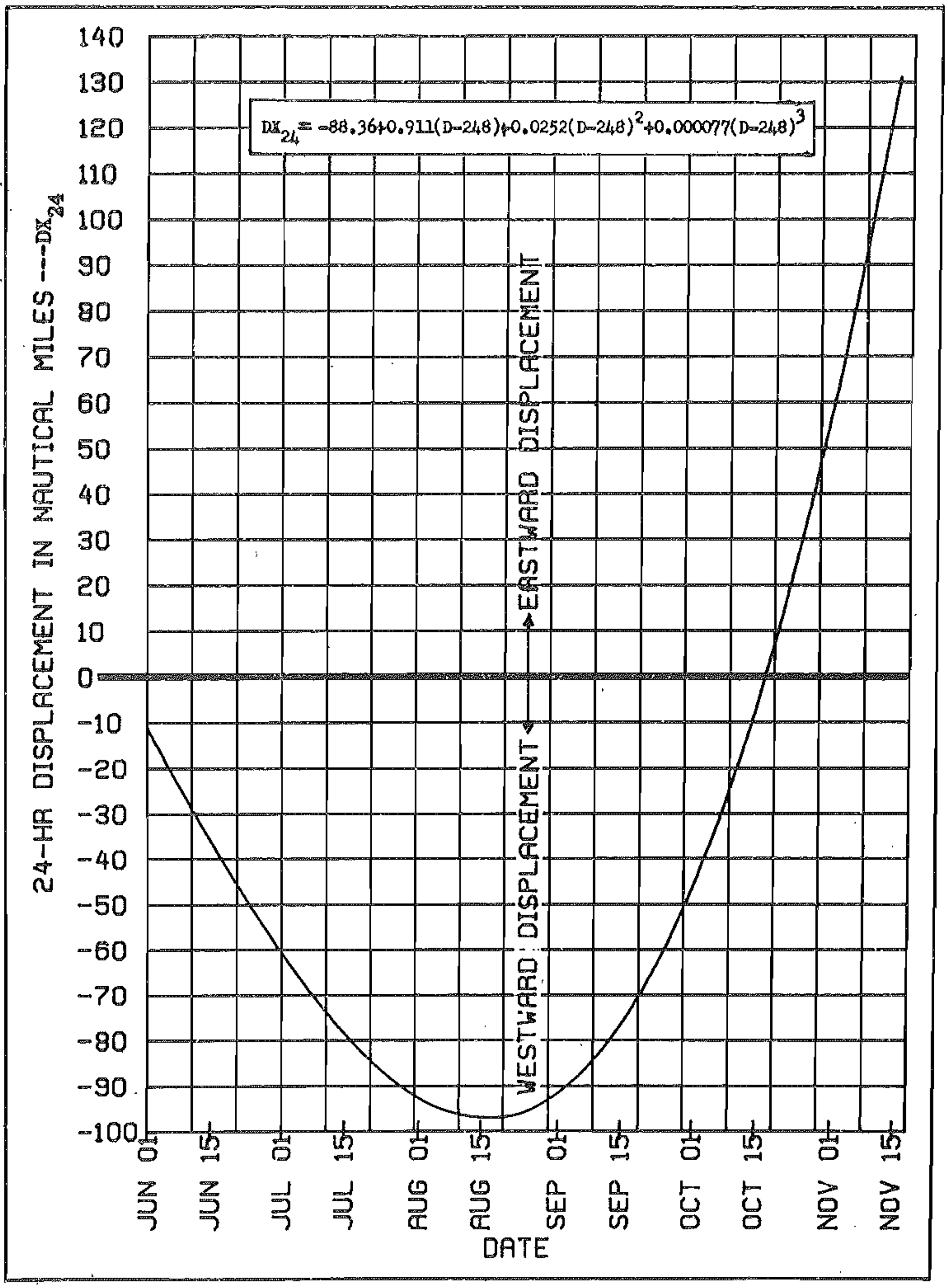

Figure 3. Twenty-four hour zonal storm displacement (DX_{24}) as a function of the third-order polynomial involving day number (D). June 1 is day number 152 and November 15 is day number 319. The multiple correlation coefficient associated with this fit is 0.22.

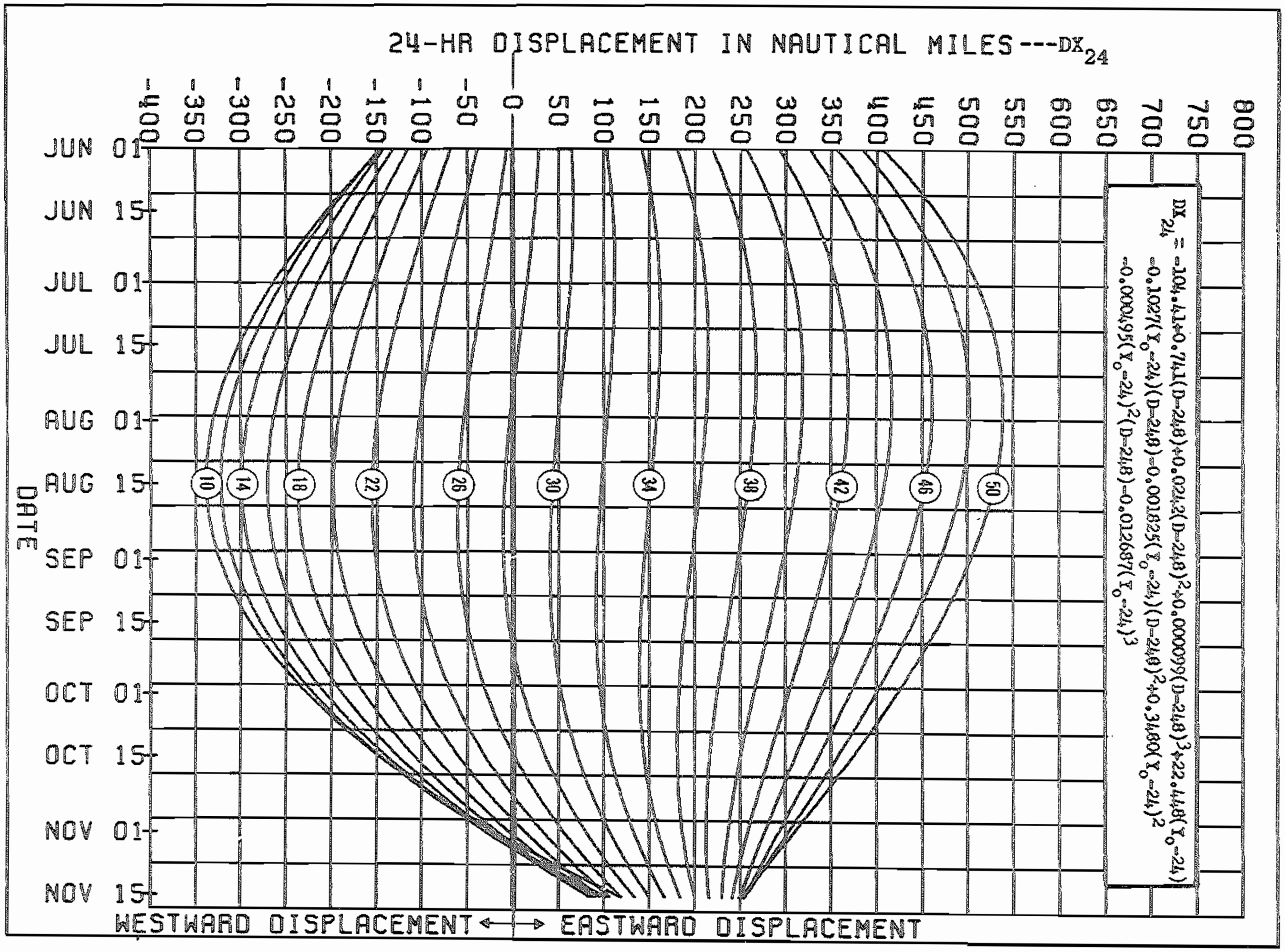

Figure 4. Twenty-four hour zonal storm displacement (DX_{24}) as a function of the third-order polynomial involving day number (D) and initial latitude (Y_o). June 1 is day number 152 and November 15 is day number 319. Circled numbers give latitude in degrees north. The multiple correlation coefficient associated with this fit is 0.71. The average latitude of storm recurvature is given by $DX_{24} = 0$.

Two further examples of non-linearity between predictand and predictors with increasing complexity in the equations are provided by figures 4 and 5. Figure 4 shows the solution to the third-order polynomial relating the 24-hr. zonal storm displacement (DX_{24}) to initial latitude (Y_o) and day number (D). Figure 5 gives the 36-hr. meridional storm displacement (DY_{36}) as given by the third-order polynomial involving latitude (Y_o), longitude (X_o) and day number (D). In each of the figures, 3, 4 and 5, the non-linear fit shows considerably greater statistical significance than does the linear fit. It is quite evident from the curvature and the non-linear gradients in the figures why this is so.

THE FINAL PREDICTION EQUATIONS

Tables 3 and 4 showed that, depending on time, between two and five predictors were selected for retention in the case of zonal motion and two through nine were retained in the case of the meridional displacement. Through considerable experimentation, it was found that a more realistic forecast track, that is, one free of unnatural discontinuities could be obtained by including all of the predictors (seven for zonal motion and thirteen for meridional motion) in each of the two sets of prediction equations. For zonal motion the six final prediction equations are given by,

$$DX_{12i} = C_{i,1} \sum_{i=1}^{6} \sum_{j=2}^{8} C_{i,j} P_j \qquad (9)$$

where the constants $C_{i,j}$ and the predictors P_j are given in Table 9. For meridional motion, the six displacement equations are given by,

$$DY_{12i} = C_{i,1} \sum_{i=1}^{6} \sum_{j=2}^{14} C_{i,j} P_j \qquad (10)$$

where the constants and predictors are listed in Table 10. Both (9) and (10) give the displacement in units of nautical miles with westward and

Table 9. Values of Constants C(I,J) for Zonal Motion

		C(I,J)					
J	P(J)	I=1	I=2	I=3	I=4	I=5	I=6
1	--- (Intercept)	-3.52591	-13.12388	-28.48156	-44.13759	-55.80913	-60.23074
2	U_0	13.69309	23.30256	32.37355	38.93667	43.27097	46.26022
3	U_{-12}	-2.63735	-3.21553	-5.34286	-6.81978	-7.86100	-8.80890
4	Y_0-24	0.81513	3.58451	8.07388	14.10797	21.27143	29.11625
5	V_0	0.68678	3.94936	9.32124	16.35476	24.07252	32.91178
6	$V_0^2U_{-12}$	-0.00217	-0.00786	-0.01318	-0.01967	-0.02254	-0.02182
7	$(Y_0-24)V_0U_{-12}$	-0.00060	-0.00676	-0.02041	-0.03853	-0.05992	-0.08554
8	X_0-68	0.12473	0.51356	1.04462	1.69802	2.47757	3.29118

southward motion considered negative. Obviously, some of the constants are too small to contribute much to the shorter range displacements. Their inclusion, however, greatly simplifies programming the equations for a digital computer.

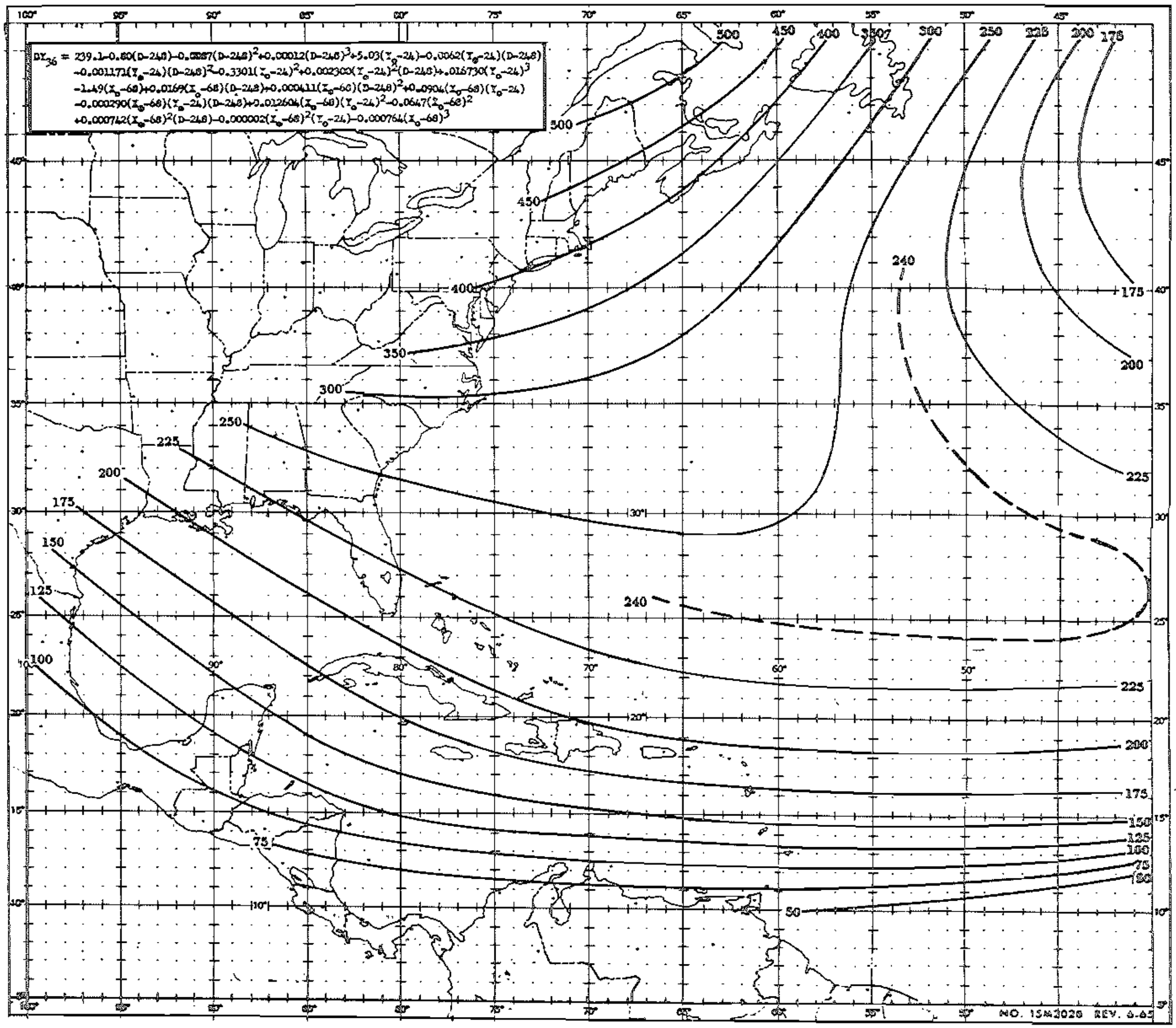

Figure 5. Thirty-six hour meridional storm displacement (DY_{36}) as a function of the third-order polynomial involving day number (D), initial latitude (Y_o) and initial longitude (X_o). In this example, the day number has been fixed at 258 (Sept. 15). The isolines give displacements northward (n. mi.) in 36 hours.

Table 10. Values of Constants C(I,J) for Meridional Motion

		C(I,J)					
J	P(J)	I=1	I=2	I=3	I=4	I=5	I=6
1	--- (Intercept)	7.60553	30.30846	67.69324	120.27143	186.02612	263.15653
2	V_0	13.59909	22.91538	31.94291	38.94701	44.48386	48.41731
3	V_{-12}	-2.57513	-2.48460	-3.69760	-4.38088	-4.72498	-4.45666
4	$V_0(V_{-12})^2$	-0.00019	0.00497	0.00967	0.01323	0.01074	0.01127
5	$(W-71)V_{-12}$	0.00460	0.00930	0.00954	0.02293	0.03200	0.04297
6	$V_0(W-71)$	0.00226	0.02511	0.06322	0.09532	0.13383	0.16962
7	$V_0{}^2V_{-12}$	-0.00149	-0.00784	-0.01332	-0.01664	-0.01607	-0.01748
8	$(Y-24)^2V_0$	-0.00027	-0.00598	-0.01611	-0.03201	-0.04866	-0.06486
9	$(D-248)^2V_{-12}$	-0.00007	-0.00035	-0.00073	-0.00122	-0.00172	-0.00222
10	$V_0(D-248)^2$	0.00004	0.00016	0.00023	0.00032	0.00036	0.00036
11	$(Y_0-24)^2(D-248)$	-0.00020	-0.00100	-0.00281	-0.00546	-0.00877	-0.01268
12	$(W-71)(D-248)V_{-12}$	0.00008	0.00048	0.00115	0.00187	0.00271	0.00369
13	U_0	0.14306	0.38795	0.89408	1.66666	2.76818	4.12125
14	$(D-248)^2$	-0.00008	-0.00067	-0.00218	-0.00435	-0.00733	-0.01102

Each of the 12 prediction equations is highly statistically significant. The maximum F-value of 9670 was associated with the 12-hour meridional displacement while the minimum F-value of 143 occurred with the 72-hour zonal equation.

Examples of CLIPER forecasts generated by equations (9) and (10) are shown in figure 6. Forecast tracks A and B represent two extremes of performance on independent data for the 1971 hurricane season. Track A (EDITH) was extremely well forecast by CLIPER. Typically, forecast systems lacking synoptic input perform well in such southerly latitudes where synoptic steering parameters show little percentage change. Track B, on the other hand, (GINGER) was poorly forecast beyond 24 hours. This is a prime example of a highly anomalous track which cannot be anticipated by the system.

Storm tracks C, D and E on figure 6 are fictitious and were selected to illustrate some additional performance characteristics of the system. Storm C is initially stationary on October 1 with a maximum wind of 50 kts. CLIPER keeps the storm center essentially stationary for the first 24 hours and provides for gradual north-northeastward acceleration thereafter. Storm tracks D show the differences when the initial maximum wind is 40 kts. (positions circled) and 100 kts. (positions indicated by squares) with otherwise similar input data on September 15. It can be noted that the zonal displacements are identical whereas the storm with the higher maximum wind has been forecast approximately 3 degrees farther north in 72 hours. Finally, storm track E illustrates how CLIPER would perform on a highly anomalous initial storm movement towards the southeast on September 15.

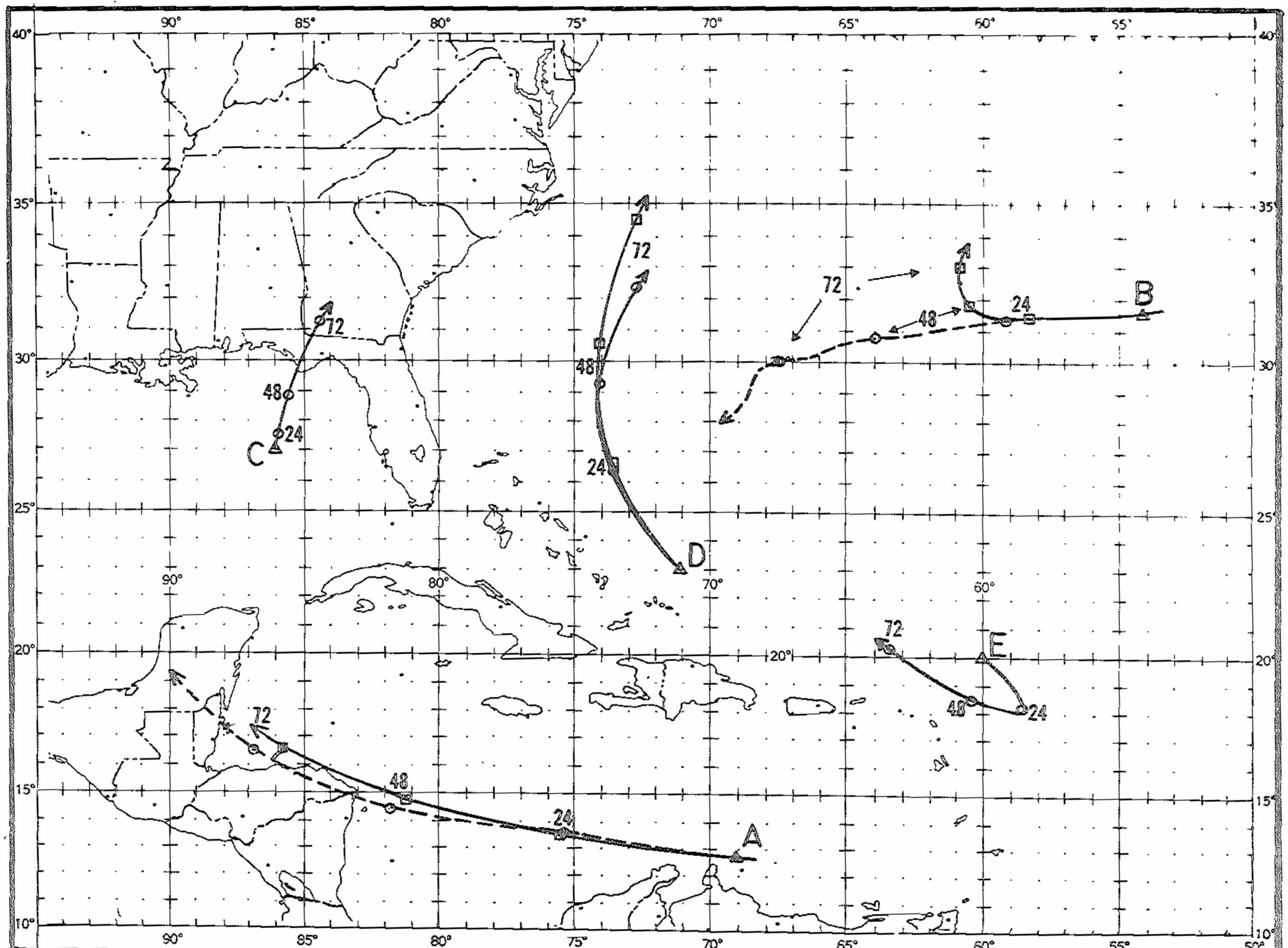

Figure 6. Example of 24, 48, and 72-hr. CLIPER forecasts (A and B), on selected cases from the 1971 hurricane season and (C,D,E), on fictitious cases discussed in text. Solid lines are CLIPER forecasts and, where applicable, dashed lines are portions of observed tracks.

ERROR ANALYSIS ON DEPENDENT DATA SAMPLE

Average errors. Tables 11 and 12 present values of the various error parameters for both the HURRAN and the CLIPER systems. Although the data in Table 11, having been extracted from Neumann and Hope, 1971, are a subset of the data contained in Table 12, the periods of record are of sufficient length to make them statistically comparable. The data have been stratified into three main groups. Type-1 storms were those having an initial westerly component of motion, type-2 an initially easterly component of motion, while type-3 includes all the cases. The rationale for this stratification was to separate those storms which had recurved from those which remained in the easterlies.

The essential difference in the errors of the two systems are illustrated in figures 7 and 8. Insofar as the mean vector errors are concerned, it can be noted that although HURRAN is somewhat better than CLIPER for storms having an initial westerly component of motion (figure 7), the two systems, on the average, give quite comparable results. However, for the storms having an initially easterly component of motion (figure 8), the CLIPER errors are significantly less than those given by the HURRAN dependent data set.

A still more definitive breakdown is afforded by plots B and C of figures 7 and 8. Here it can be noted that greater differences between the two systems occur in the case of 36 and 48-hour meridional motion of type-2 storms with the CLIPER errors clearly less than the HURRAN errors. The latter outperforms CLIPER only in the case of the zonal component of storm motion prior to recurvature (figure 7,B). It should be noted however, that the type-1 storms are usually the type which affect land areas over the Caribbean and the Continental United States.

Errors as a function of initial storm motion. As suggested by figures 7 and 8 and as pointed out by Neumann and Hope, 1971. the vector errors (E) are a function of the initial orthogonal components of storm motion (U_o,V_o) such that,

$$E = f(U_o,V_o). \qquad (11)$$

Representing this function by a third-order polynomial,

$$E(U_o,V_o) = \sum_{i,j} C_{ij} U_o^{\,i} V_o^{\,j} \qquad (i+j \leq 3) \qquad (12)$$

gives an expression which, when expanded,yields ten terms and 10 constants, C . Ten normal equations are required to solve for these constants by least squares techniques. Figure 9 shows the graphical solution of the resultant regression surface for the 36-hour vector errors. The elliptical boundary to the analysis represents the 99 percent data envelope assuming the joint distribution of U_o and V_o is bivariate normal. Ezekiel, 1941, points out that serious errors can be

Table 11. Summary of HURRAN performance on data sample extending from 1945 through 1969.

	Type 1 storms (Initial westward component)					Type 2 storms (Initial eastward component)					Type 3 storms (All cases)				
Forecast period (hr)	12	24	36	48	72	12	24	36	48	72	12	24	36	48	72
Number of cases	403	401	382	367	317	225	220	200	164	77	628	621	582	531	394
Mean vector error (n.mi.)	24	62	115	180	327	40	125	241	354	556	30	84	158	234	372
Standard deviation of errors (n.mi.)	15	37	69	112	226	25	74	132	187	304	21	61	113	161	259
Mean U component error (longitude) bias* (n.mi.)	-1	-1	0	-2	9	-1	-5	-25	-25	-46	-1	-2	-9	-9	-2
Mean V component error (latitude) bias* (n.mi.)	-2	5	7	8	0	3	0	-5	10	89	2	4	3	9	17
Standard deviation of U component error (n.mi.)	21	51	95	151	284	34	103	193	282	504	27	74	137	201	338
Standard deviation of V component error (n.mi.)	19	51	95	149	279	33	109	195	283	376	25	73	138	200	302
Mean Absolute value U component error (n.mi.)	15	39	73	118	219	26	80	152	225	410	19	54	100	151	256
Mean Absolute value V component error (n.mi.)	15	41	76	116	210	26	81	155	227	316	19	55	103	150	231

Table 12. Summary of OLIPER performance on data sample extending from 1931 through 1970.

	Type 1 storms (Initial westward component)					Type 2 storms (Initial eastward component)					Type 3 storms (All cases)				
Forecast period (hr)	12	24	36	48	72	12	24	36	48	72	12	24	36	48	72
Number of cases	2130	2130	2130	2130	2130	1026	1026	1026	1026	1026	3156	3156	3156	3156	3156
Mean vector error (n.mi.)	22	70	124	190	341	30	103	190	290	508	25	80	146	222	395
Standard deviation of errors (n.mi.)	16	49	84	124	222	21	75	126	178	286	18	61	104	151	257
Mean U component error (longitude) bias* (n.mi.)	1	4	6	9	11	-2	-8	-13	-18	-23	0	0	0	0	0
Mean V component error (latitude) bias* (n.mi.)	1	1	3	5	12	-2	-3	-6	-11	-25	9	0	0	0	0
Standard deviation of U component error (n.mi.)	21	65	116	177	319	26	94	171	257	446	23	76	136	207	366
Standard deviation of V component error (n.mi.)	17	55	95	141	252	26	86	150	221	374	21	66	116	172	298
Mean Absolute value U component error (n.mi.)	15	49	88	135	245	19	68	129	200	359	16	55	101	156	282
Mean Absolute value V component error (n.mi.)	12	40	70	107	192	20	63	115	173	294	15	47	85	128	225

*Forecast minus observed.

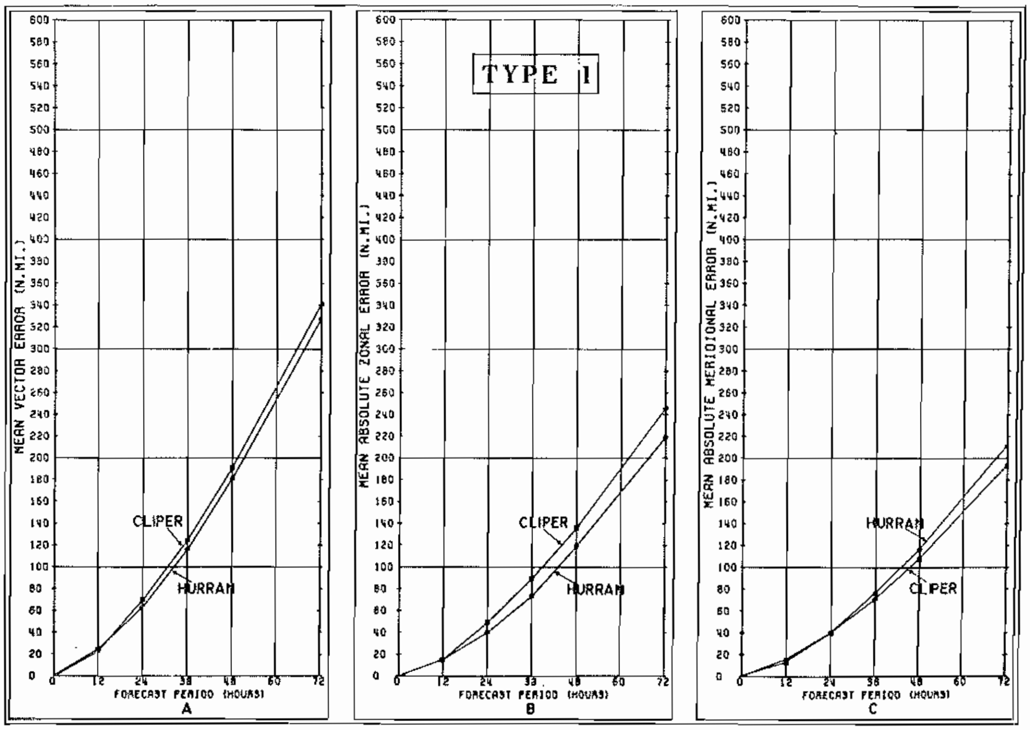

Figure 7. Dependent data forecast errors (n.mi.) for the HURRAN and CLIPER systems for type 1 (initial westerly component) storms.

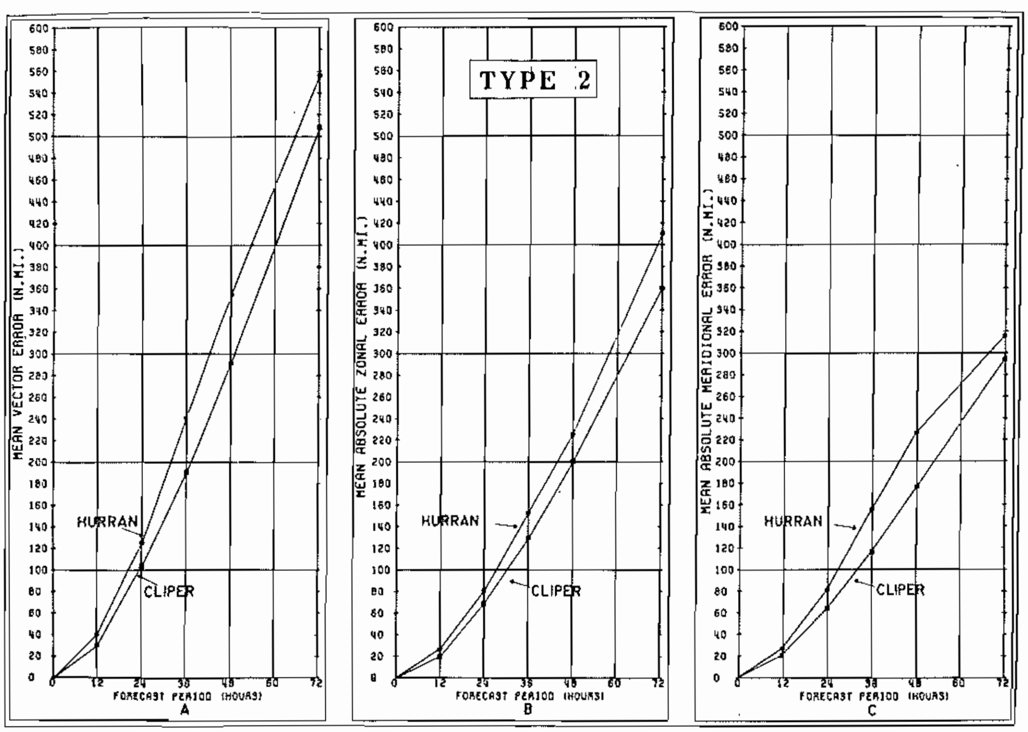

Figure 8. Dependent data forecast errors (n.mi.) for the HURRAN and CLIPER systems for type 2 (initial easterly component) storms.

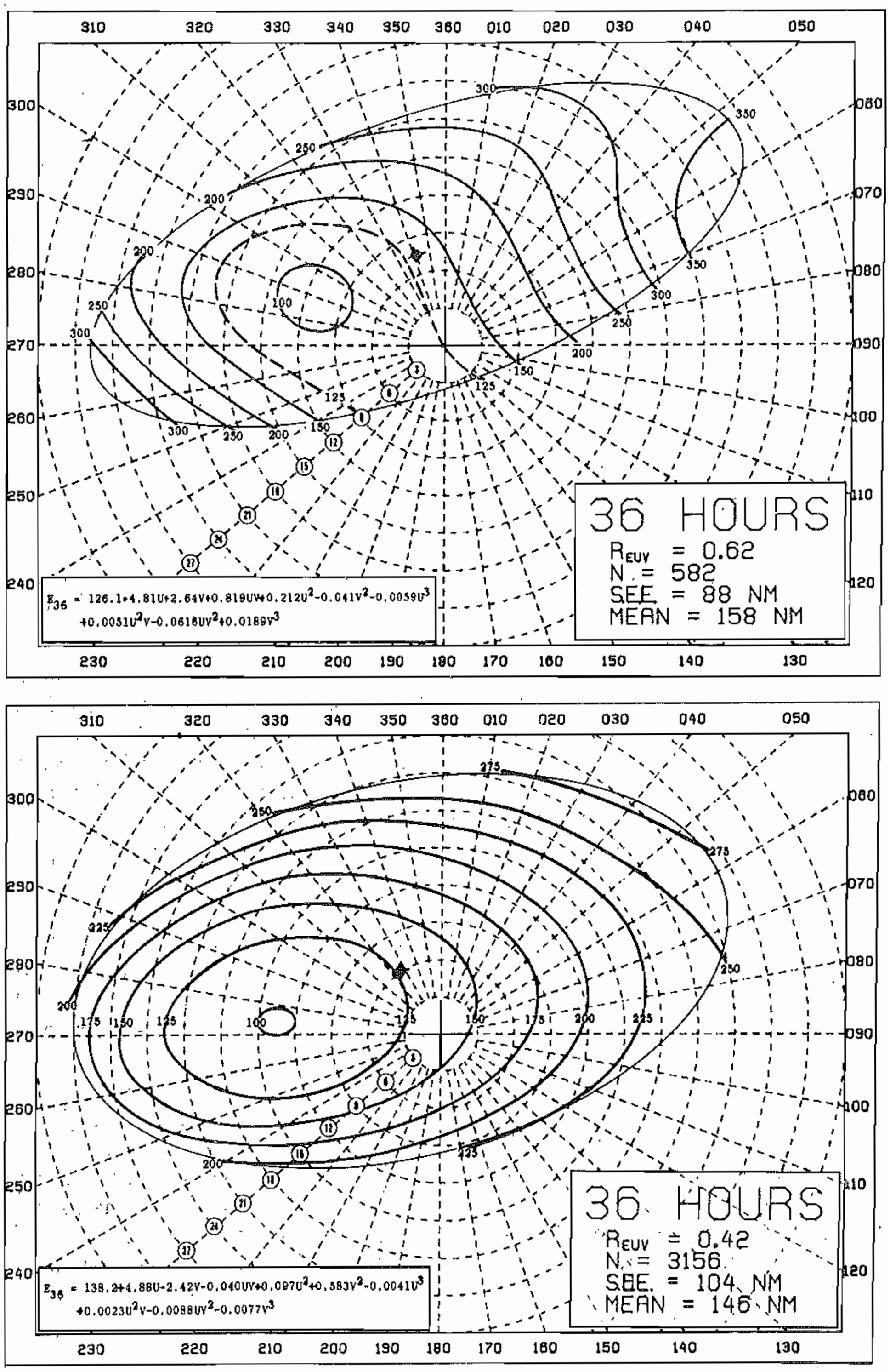

Figure 9. Thirty-six hour HURRAN (Top) and CLIPER (Bottom) errors in n.mi. (solid curves) as a function of the initial storm heading and speed. The dashed radials are storm headings in degrees and the concentric dashed circles are spaced at 3 kt. intervals. The ellipse is the 99 percent data envelope. R_{EUV} is multiple correlation coefficient, N is number of cases and S.E.E. is standard error.

realized by attempting to apply such a regression function outside the bounds of the dependent data sample; hence, the elliptical constraint.

To use figure 9, one enters the polar diagram with the initial speed and direction of the storm. With an initial storm motion of 300/09, for example, the HURRAN system can be expected to yield a 36-hour vector error of 100 n.mi. while CLIPER gives about 115 n.mi. Figure 9 points out the exact areas over which each system can be expected to statistically excel.

RESTRICTIONS ON THE USE OF CLIPER

The CLIPER system provides the forecaster with an estimate of tropical cyclone displacement based on a least-squares fit to predictors or predictor functions derived from a storm's initial motion, its past motion, the present location, the time of year and the storm intensity. Since synoptic parameters which influence storm motion are not included, the forecaster must be prepared to make subjective corrections to the displacements whenever warranted by anomalous pressure-patterns. Objective techniques to make such corrections are under development and will be reported on in future papers.

Obviously, the errors of the system will be greater when using independent data. This is not so much that the development equations are tailored to the dependent data but rather that the forecaster does not always know the current and past storm motion as well as might be desired. Every effort should be made to define these input parameters as precisely as possible. Only in this way will the full variance reducing potential of the system be realized.

Figure 1 showed the spatial distribution of the dependent data set. Occasionally, a storm may be observed outside the bounds of this set. In this event, use of the CLIPER equations may well give erratic displacements. Other statistical regression schemes, of course, are subject to this same restriction.

The statistical significance F-test based on the 3156 sets of dependent data, the correlation coefficient and the loss of degrees of freedom showed that the prediction equations were statistically significant at well below the 1 percent level. However, the 3156 sets of data are not entirely independent, effectively lowering the number of cases. But, even reducing the number of observations by 50% still gives sufficient statistical significance.

CONCLUSIONS

Development of the CLIPER equations and the resultant error analysis lead to the following conclusions:

1. Although the CLIPER system, as originally conceived, was intended as an alternate to HURRAN, its comparative simplicity and relatively low residual errors justify its use as an independent objective technique.

2. The error analysis shows that CLIPER outperforms HURRAN on storms which have recurved into the westerlies. Although HURRAN does better than CLIPER on storms which remain in the easterlies, the difference is quite small.

3. Considerably more variance is explained by the CLIPER system for zonal motion than for meridional motion. This suggests that the addition of synoptic predictors should improve the latter more than the former.

4. The primary predictor for all displacement equations is the current meridional or zonal component of storm motion. The forecaster should make every effort to insure that the current (and past 12-hour) storm motion is described as precisely as the data permit.

LIST OF REFERENCES

Akima, H., 1970: "A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures", Journal of the Association For Computing Machinery, Vol. 17, No. 4 pp. 589-602.

Burington, R.S. and D.C. May, 1958: Handbook of Probability and Statistics, Handbook Publishers, Inc., Sandusky, Ohio, pp. 276-277.

Cry, G.W., W.H. Haggard, and H.S. White, 1959: "North Atlantic Tropical Cyclones", Technical Paper 36, U. S. Department of Commerce, Weather Bureau, Washington, D. C., 214 pp.

Efroymson, M.A., 1964: "Multiple Regression Analysis", Mathematical Methods for Digital Computers, edited by A. Ralston and H. S. Wilf, John Wiley and Sons, Inc., New York, pp. 191-203.

Ezekiel, M., 1941: Methods of Correlation Analysis, Wiley and Sons, New York, pp. 347-349.

Hope, J. R., 1971: "Position Error of Tropical Cyclones Vs. Frequency of Aircraft Fixes", Journal of Applied Meteorology, Vol. 10, No. 5, pp. 921-936.

Hope, J.R. and C.J. Neumann, 1968: "Probability of Tropical Cyclone Induced Winds at Cape Kennedy", ESSA Technical Memorandum WBTM SOS-1, 67 pp.

Hope, J.R. and C.J. Neumann, 1970: "An Operational Technique for Relating the Movement of Existing Tropical Cyclones to Past Tracks", Monthly Weather Review, Vol. 98, No. 12, pp. 925-933.

Miller, B.I. and P.P. Chase, 1966: "Prediction of Hurricane Motion by Statistical Methods", Monthly Weather Review, Vol. 94, No. 6, pp. 389-406.

Miller, B.I., E.C. Hill and P.P. Chase, 1968: "Revised Technique for Forecasting Hurricane Motion by Statistical Methods", Monthly Weather Review, Vol. 96, No. 8, pp. 540-548.

Mills, F.C., 1955: Statistical Methods, Holt, Rinehart and Winston, New York, pp. 631-656.

Neumann, C.J. and J.R. Hope, 1971: "A Performance Analysis of the HURRAN Tropical Cyclone and Forecast System" (submitted to MWR for publication).

Panofsky, H.A., 1949: "Objective Weather-Map Analysis", Journal of Meteorology, Vol. 6, No. 6, pp. 386-392.

Sanders, F. and R. W. Burpee, 1968: "Experiments in Barotropic Hurricane Track Forecasting", Journal of Applied Meteorology, Vol. 7, No. 3, pp. 313-323.

Simpson, R.H., 1971: "The Decision Process in Hurricane Forecasting", NOAA Technical Memorandum NWS SR-53, 35 pp.